



PATENT
ASSIGNEE: INTEL Corporation

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appl. No. : 09/397,850 Confirmation No. 9295
Applicant : Arlin R. Davis
Filed : September 17, 1999
Group Art Unit : 2157
Examiner : Barbara N. Burgess
For : METHOD OF READING A REMOTE MEMORY

Docket No. : 02207/7240
Customer No. : 23838

RECEIVED

FEB 20 2004

Technology Center 2100

COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION PURSUANT TO 37 C.F.R. 1.131

I, Arlin R. Davis, hereby declare the following:

1. I am the sole inventor of the subject matter claimed in U.S. Patent Application Serial No. 09/397,850, filed September 17, 1999 and entitled "METHOD OF READING A REMOTE MEMORY."

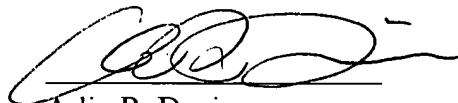
2. The invention described and claimed in the present application was conceived prior to July 29, 1999. Evidence of this fact is shown in the invention disclosure form attached as Exhibit A hereto, which was prepared and submitted to my employer at the time of the disclosure, Intel Corporation, prior to July 29, 1999. In addition, the invention described and claimed in the present application was also completed by actual reduction to practice prior to July 29, 1999.

3. I exercised diligence in, at a minimum, constructively reducing the claimed invention to practice from at least a time prior to July 29, 1999 continuously up to September 17, 1999, the date on which the above-cited non-provisional patent application was filed. During that time, I provided information to patent counsel for preparation of the application, and reviewed/revised drafts of the application that was filed on September 17, 1999.

I, Arlin R. Davis, acknowledge that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. § 1001) and may jeopardize the validity of the above-cited non-provisional patent application or any patent issuing thereon. Likewise, I declare under penalty of perjury that the above statements are true and correct to the best of my knowledge, information and belief.

Respectfully submitted,

Dated: 2-10-2004

A handwritten signature in dark ink, appearing to read 'Arlin R. Davis', written over a horizontal line.

Arlin R. Davis

Arlin R. Davis

INTEL INVENTION DISCLOSURE

LEGAL ID#

11348Platforms
Comm

ESG/NGIO

DATE: _____

It is important to provide accurate and detailed information on this form. The information will be used to evaluate your invention for possible filing as a patent application. When completed, please return this form to the Legal Department at JF3-147. If you have any questions, please call 264-0444 or 264-0998.

1. ✓ Inventor: Davis Arin R
 Last Name First Name Middle Initial
 SS# 225-92-7991 WWID 10030424 Phone (503) 677-7748 M/S: CO3-202
 Home Address: 23150 NE Cove Orchard City Yamhill State OR Zip 97148
 Citizenship: USA BUM Presenter: Tom MacDonald
 Group: (e.g. TMG, ICG, CEG) ESG Division Name Next Generation I/O Subdivision _____
 Supervisor * Mark Wunderlich WWID 10039305 Phone (503) 677-5567 M/S: CO5-201

(PROVIDE SAME INFORMATION AS ABOVE FOR EACH ADDITIONAL INVENTOR)

- ✓ 2. Title of Invention: Method of Reading Remote Memory across a Virtual Interface
3. What technology/product/process (code name) does it relate to: NGIO and VI Architecture
4. Stage of development (i.e. % complete) 100%
5. (a) Has a description of your invention been, or will it shortly be, published outside Intel:
 NO: _____ YES: X DATE WAS OR WILL BE PUBLISHED: _____
 If YES, was the manuscript submitted for pre-publication approval? YES: _____ NO: X
- (b) Has your invention been used/sold or planned to be used/sold by Intel or others?
 NO: X YES: _____ DATE WAS OR WILL BE SOLD: _____
- (c) Does this invention relate to technology that is or will be covered by a SIG (special interest group)/standard/ or specification?
 NO: _____ YES: X Name of SIG/Standard/Specification: NGIO Architecture: Master Specification and VI Architecture Specification
- (d) If the invention is a semiconductor device, actual or anticipated date of tapeout? _____
- (e) If the invention is software, actual or anticipated date of any beta tests. _____
6. Was the invention conceived or constructed in collaboration with anyone other than an Intel blue badge employee or in performance of a project involving entities other than Intel, e.g. government, other companies, universities or consortia?
 NO: X YES: _____ Name of individual or entity: _____

PLEASE READ AND FOLLOW THE DIRECTIONS ON THE ATTACHED
 PAGE ON HOW TO WRITE A DESCRIPTION OF YOUR INVENTION

RECEIVED

FEB 20 2004

Technology Center 210

Please attach a page to this form, DATED AND SIGNED BY AT LEAST ONE PERSON WHO IS NOT A NAMED INVENTOR, to provide a description of the invention, and include the following information:

1. Describe in detail how the invention works.
2. Describe advantage(s) of your invention over what is done now.
3. Include at least one figure illustrating the invention. If the invention relates to software, include a flowchart or pseudo-code representation of the algorithm.
4. Value of your invention to Intel (how will it be used?).
5. Identify the closest or most pertinent prior art that you are aware of.
6. Who is likely to want to use this invention or infringe the patent if one is obtained and how would infringement be detected?

*HAVE YOUR SUPERVISOR READ, DATE AND SIGN COMPLETED FORM

DATE: _____

SUPERVISOR: *Mark W. Wundt*

BY THIS SIGNING, I (SUPERVISOR) ACKNOWLEDGE THAT I HAVE READ AND UNDERSTAND THIS DISCLOSURE, AND RECOMMEND THAT THE HONORARIUM BE PAID

Method of Reading Remote Memory across a Virtual Interface

Purpose:

This invention describes a method by which remote direct memory read access can be performed across a Virtual Interface (VI) that does not provide such capabilities. The method employs a simple transport protocol between the host-side requesting I/O data transfers and the device-side performing the I/O data transfer operations. The Virtual Interface's remote Direct Memory Access (rDMA) write and immediate data services are utilized by the host-side to provide the necessary I/O data transfer and completion notification.

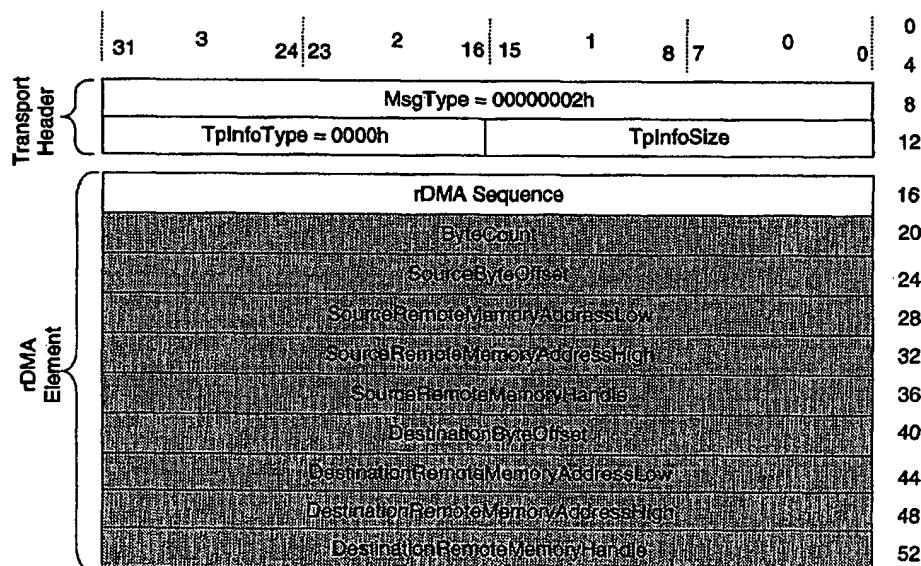
Description:

Optimal rDMA read capabilities requires both a well-defined communication interface, such as that defined by the Virtual Interface Architecture (VIA), along with a new method of rDMA read request processing. The remote read processing is performed with new software techniques that are described in detail below.

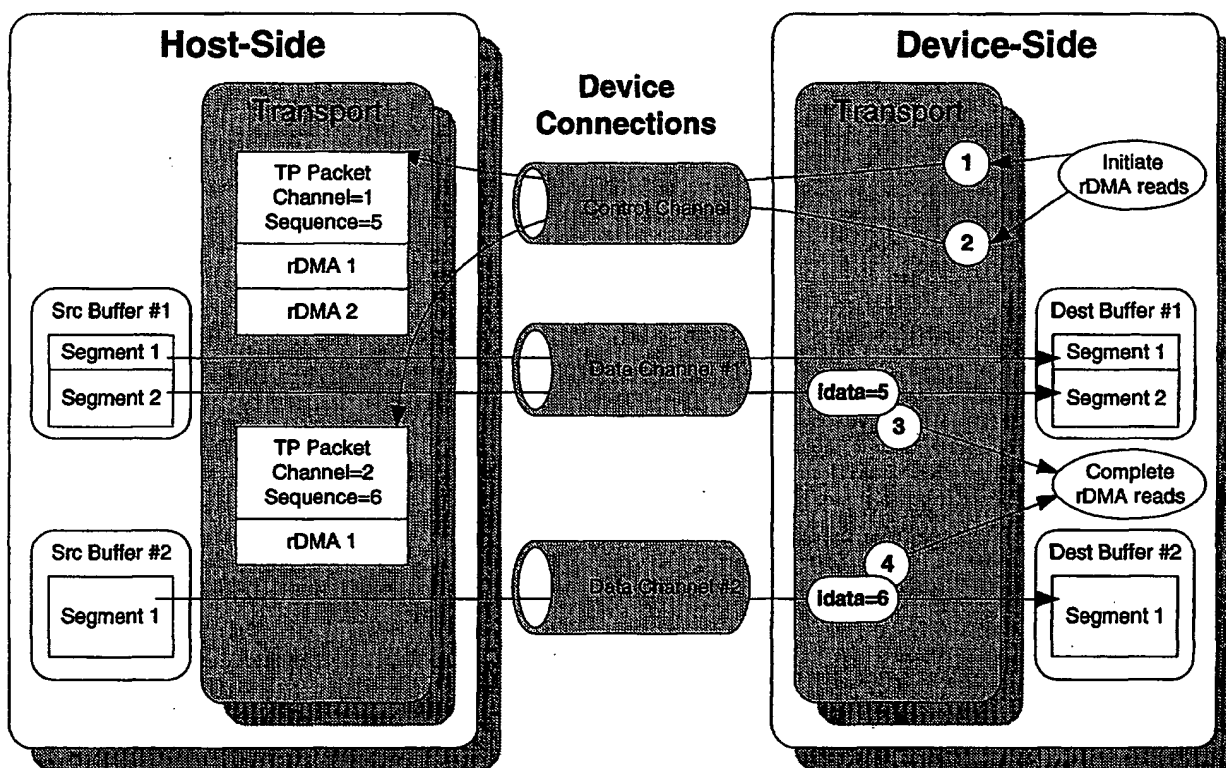
Remote DMA Read Initiation from the Device-side

To initiate a rDMA read request, a device-side driver provides a set of source buffers, which reference the remote host-side memory, and a set of destination buffers that reference the local device-side memory. These buffers are pre-registered with their respective VI network interface controller (NIC) before the read request is initiated. The remote read mechanism allows these buffers to be specified as multiple data segments with offsets and a VI channel designation, which indicates the data path. The rDMA Sequence field is used to specify the last data segment and completion of the rDMA read request.

The rDMA read request interface describes the data transfer from the host source buffers to the device destination buffers by using rDMA elements. The rDMA elements are embedded in a lightweight transport message that is identified by a unique header type and is recognized by the host-side as a rDMA read request. These rDMA elements along with the transport header, which specify both the type (TpInfoType) and count (TpInfoSize), are illustrated by the figure below.



After receiving the rDMA read request, the host-side will schedule rDMA write transactions that are directed back to the device-side in order to complete the request. This rDMA read request is illustrated by the figure below.



Two separate initiated rDMA read requests are shown in the figure above to help illustrate the design of the new rDMA read method. This example highlights the use of multiple segments to support scatter/gather and the use of multiple data channels that multiplex separate rDMA read transfers. The first rDMA read request (step 1) is scheduled for Data Channel 1 and the second rDMA read request (step 2) is scheduled for Data Channel 2 to avoid blocking and optimize performance. The rDMA read mechanism identifies each rDMA read with a unique sequence that will be used to complete the read request. The rDMA read mechanism on the host-side recognizes the incoming packets on the control channel as embedded rDMA read requests and builds the proper VI rDMA write descriptors with the proper sequence inserted into the immediate data field on the last segment of each rDMA read request. The rDMA read completion for both requests are processed on the device-side, based on the immediate data that arrives with the last data segments (steps 3 and 4) of each rDMA transfer. This immediate data eliminates the need for the host-side to send a data transfer acknowledgement packet back to the device-side.

Advantages of the rDMA Read Technique:

- ◆ Provides an I/O communication mechanism that allows a remote attached I/O device to read remote host memory with minimal host-side CPU involvement. Typical Virtual Interface network interface controllers do not support rDMA read methods.
- ◆ Eliminates the need to acknowledge the completion of the rDMA write by using immediate data services provided by VI. Since the rDMA completion is identified with a sequence and end of data context, the receiving I/O unit can immediately identify the completing rDMA read request without waiting for an acknowledgement packet from the host. This method does not only eliminate an acknowledgement packet on the wire but also provides a very fast completion mechanism on the receiving I/O unit.
- ◆ Allows multiplexed rDMA read transactions to occur across multiple VI channels to provide optimum performance and eliminate possible blocking conditions. This method gives the I/O device the flexibility to separate the large and small data transfers so that small transfers will not be blocked behind large transfers.
- ◆ Provides the capability for multiple data segments and strict in-order delivery across each VI channel. Multiple data segments provided by this technique provide devices with an efficient mechanism to support a typical scatter/gather list of data buffers. In-order scheduling of the rDMA read across each channel provides a guarantee of data transfer ordering.

Benefit to Intel:

This invention strengthens Intel's position for enhanced server solutions by providing software mechanisms that are required for remote I/O device processing. In addition, this capability allows Intel and it's collaborators the ability to emulate a NGIO communication link which in turn enable software stack development on both hosts and IO units ahead of NGIO hardware availability.

Infringement information:

The capability that this invention provides could result in the availability of remote I/O devices based on VI architecture, which do not currently exist. Additional server capabilities along with better performance and scalability of remote I/O devices brings about product differentiation in the market place. These advantages may be advertised by OS vendors or I/O vendors, which in turn would make the patent visible.

Submitted by: Arlin Davis

Witnessed by: Mark Wunderlich,  Dated: